Supplementary Materials for Enforcing Monotonous Shape Growth or Shrinkage in Video Segmentation

Yuliya Tarabalka
yuliya.tarabalka@inria.fr
Inria Sophia-Antipolis, AYIN team
Sophia Antipolis, France

Guillaume Charpiat
guillaume.charpiat@inria.fr
Inria Sophia-Antipolis, STARS team

Ludovic Brucker
ludovic.brucker@nasa.gov
NASA GSFC & Universities Space Research Association
MD, USA

Bjoern H. Menze
bjoern@ethz.ch
Computer Vision Laboratory
ETH Zurich, Switzerland &
Inria Sophia-Antipolis, Asclepios team

1 Details about section 2.4 : Rewriting as a multi-label problem

We show now another point of view on sequence segmentation with growth constraint. The successive labels $L_i(t)$ of a given pixel i over time might change only once, and only from 0 (background) to 1 (foreground object). Hence, this vector of labels $L_i(t)$ is of the form $(0,0,\ldots,0,1,\ldots,1,1)$ and can be represented by just the time index τ_i of the first 1, *i.e.* the earliest time at which the pixel starts belonging to the object. This time τ_i is in $[1,T+1]$, with $T+1$ meaning "never". We have thus transformed a binary optimization problem on a sequence of images with shape growth constraint into a multi-label problem defined on one single image, without any constraint. This new problem can be expressed in the Markov Random Field form (1) with $V_i(\tau_i) := \sum_{t<\tau_i} V_i^t(0) + \sum_{t\geqslant\tau_i} V_i^t(1)$ and

$$W_{i,j}(\tau_i,\tau_j) := \sum_{t<\min(\tau_i,\tau_j)} W_{i,j}^t(0,0) + \sum_{\tau_i\leqslant t<\tau_j} W_{i,j}^t(1,0) + \sum_{\tau_j\leqslant t<\tau_i} W_{i,j}^t(0,1) + \sum_{t\geqslant\max(\tau_i,\tau_j)} W_{i,j}^t(1,1)$$

where (V_i^t) and $(W_{i,j}^t)$ define the energy E^t at time t, and where sets of summation may be empty. The submodularity of the binary interaction terms W^t in each frame implies the submodularity of the multilabel interaction term W, *i.e.*

$$W_{i,j}(\tau_1,\tau_2) + W_{i,j}(\tau_1',\tau_2') \leqslant W_{i,j}(\tau_1,\tau_2') + W_{i,j}(\tau_1',\tau_2)$$

for all labels satisfying $\tau_1 \leqslant \tau_1'$ and $\tau_2 \leqslant \tau_2'$. Thus this energy can be minimized globally efficiently with now standard techniques (*e.g.* [1]). Note that in the particular case where interaction terms W^t do not depend on t, the interaction term W of this multi-label energy above can be rewritten as a convex function g of $(\tau_i - \tau_j)$, and then Ishikawa's construction [2] can be applied. It turns out that the graph built this way is precisely the graph that we built in our initial binary multi-frame problem. Our initial formulation is however more flexible, in that interaction terms can depend on t, and more natural, in that inclusion constraints can easily be enforced in spatial or/and time subregions only, while this would not be expressible with the multi-label formulation.

References

[1] J. Darbon. Global optimization for first order Markov random fields with submodular priors. In *Int. Workshop on Combinatorial Image Analysis*, 2008.

[2] H. Ishikawa. Exact optimization for Markov random fields with convex priors. *TPAMI*, 25, October 2003.